Modeling and Prediction of Online Product Review Helpfulness: A Survey

Gerardo Ocampo Diaz and Vincent Ng | The University of Texas at Dallas

Introduction

User reviews posted on e-commerce and review websites vary in quality and are too many for customers to read. There is a growing need for systems that can automatically assess the helpfulness of reviews.

Our Goals

- Provide an overview of the most relevant work in the last decade
- Discuss gained insights and provide guidelines for future research

Datasets

1. Amazon Review Dataset (ARD) (McAuley et al., 2015)
2. Multi-Domain Sentiment Dataset (MDSD) (Blitzer et al., 2007)
3. Ciao Dataset (Tang et al., 2013)

ASIN	X	X	–
Product Name&Category	X	X	X
Product Metadata	–	X	–
Star Rating	X	X	X
Review Title	X	X	–
Review Text & Post Time	X	X	X
Reviewer Name	X	X	–
Reviewer ID	–	X	X
Helpfulness Vote Ratio	X	X	–
Vote - User ID Pairs	–	–	X
Social Network	–	–	X
	MDSD	ARD	CIAO

Helpfulness Tasks

- **Regression**
- **Binary Classification**
- **Ranking**

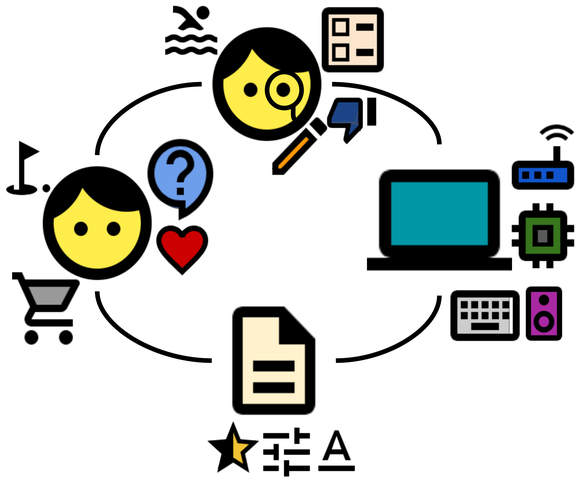
Features

- **Content Features**
  - Review length (words), star rating, readability metrics, UGR TF-IDF, sentence subjectivity, word lists.
- **Context Features**
  - Reviewer experience., avg. helpfulness, user - reviewer connections, user - reviewer product rating similarity.

State of Helpfulness Prediction

- Strong focus on new features.
- Advances hindered by the lack of **standard datasets**, **defined baselines** and **feature ablation**.
- NN approaches have shown performance increases with (Malik and Hussain, 2017) and without handcrafted features (Chen et al., 2018); also, domain knowledge transfer (Chen et al., 2018).
- **User-specific systems** have shown performance increases (Tang et al., 2013).
- Semantic-focused models have shown increased performance (Malik and Hussain, 2017).

Moderating Factors



- A closer look at the **entities** involved in the **voting process** and **interactions** between them has led to the discovery of **moderating factors** i.e., mechanisms and properties that can influence the voting process outcome.
- **User-Product Predispositions**
  - Difference of user and reviewer opinions on the product..
- **User-Reviewer Idiosyncrasy**
  - Product rating similarity and connections in social network.
- **Product Nature**
  - Search vs. Experience products.
- **Review Nature**
  - Critical, comparative, suggestive?
- **Review Context**
  - Relative helpfulness compared to neighboring reviews.

Conclusions

- A multi-faceted, contextual task.
- A lack of standard datasets, baselines, feature ablation.
- Understanding the voting process is necessary.
- Promising systems focus on user-specific helpfulness and feature interactions.

Recommendations

- **Task**: Focus on user-specific helpfulness prediction.
- **Dataset**: Create a standard that facilitates the design of user-specific models.
- **Knowledge Sources**: Inspiration from recommendation systems: user profiling and review, user "similarity".
- **Baseline**: Study feature ablation, enhance "leading" systems: Tang et al. (2013), Mukherjee et al. (2017), Malik and Hussain (2017), Chen et al. (2018).

References

1. John Blitzer, Mark Dredze, and Fernando Pereira. 2007. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 440–447.
2. Cen Chen, Yinfei Yang, Jun Zhou, Xiaolong Li, and Forrest Sheng Bao. 2018. Cross-domain review helpfulness prediction based on convolutional neural networks with auxiliary domain discriminators. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 602–607.
3. M.S.I. Malik and Ayyaz Hussain. 2017. Helpfulness of product reviews as a function of discrete positive and negative emotions. Computers in *Human Behavior*, 73:290–302.
4. Julian McAuley and Jure Leskovec. 2013. Hidden factors and hidden topics: Understanding rating dimensions with review text. In *Proceedings of the 7th ACM Conference on Recommender Systems*, pages 165–172.
5. Subhabrata Mukherjee, Kashyap Popat, and Gerhard Weikum. 2017. Exploring latent semantic factors to find useful product reviews. In *Proceedings of the 2017 SIAM International Conference on Data Mining*, pages 480–488.
6. Jiliang Tang, Huiji Gao, Xia Hu, and Huan Liu. 2013. Context-aware review helpfulness rating prediction. *In Proceedings of the 7th ACM Conference on Recommender Systems*, pages 1–8.

ARD: http://jmcauley.ucsd.edu/data/amazon *MDSD:* https://www.cs.jhu.edu/~mdredze/datasets/sentiment/ *CIAO:* https://www.cse.msu.edu/~tangjili/trust.html

THE UNIVERSITY OF TEXAS AT DALLAS
Erik Jonsson School of Engineering and Computer Science